How a Type-Type Identity Theorist Can Be a Non-Reductionist: An Answer from the Idealizational Conception of Science

KATARZYNA PAPRZYCKA

Two constraints are at work in our philosophical conceptions of the mental. First, we think that the mental is somehow radically different from the physical. However, the cost of underwriting this intuition ontologically is the mind-body problem, whose solution is monism. This has in part motivated the second intuition, viz. that the world is ontologically homogenous—it is made of just one kind of stuff, matter. The most natural form of materialism is the type-type identity theory, according to which there is an identity not only between mental and physical event-tokens but also between mental and physical event-types. It has been widely agreed that this version of monism straightforwardly leads to reductionism, the conviction that higher-level sciences (in particular psychology) will be in the end reduced to physics. This in turn leads to a straightforward denial of the first intuition, that the mental is different from the physical.

One way of trying to reconcile both intuitions has been the position of non-reductive materialism. One of the best known proposals is Davidson's anomalous monism, one of whose distinctive claims is that there are no type-type but only token-token identities among the mental and physical events.[1] The reception of the token-token identity theory has been mixed. On the one hand, it provides a solid basis for the position of anti-reductionism and thus for reconciling both intuitions. On the other hand, however, it has been faced with the objection of type-epiphenomenalism, according to which if one gives up type-type identities, it becomes unclear how the mental as mental can be viewed as being causally efficacious at all [Soutland, 1976; Stoutland, 1985; Honderich, 1982; McLaughlin,

[1] I will not be concerned here with Davidson's second distinctive claim, which actually motivates his version of token-token identity theory. Davidson believes that there are no psychological (or psychophysical) laws in part because folk-psychological generalizations involve content-bearing propositional attitudes. Davidson provides no reasons, however, to believe that the science of psychology will be modelled on folk psychology. His paradigm case of psychological research is the study of the limitations of formal decision theory as a psychological theory, which seems rather special and not at all paradigmatic of the science of psychology as it stands, not to mention any future psychology.

Laws and Models in Science, 113–128.

1989; Kim, 1993a].

It is clear that the issue at stake lies on the borderline between philosophy of mind, metaphysics and philosophy of science. Most of the current debate is carried out among philosophers of mind and metaphysicians. The aim of this paper is to show how the debate can be enriched by drawing on some developments in philosophy of science. I will show in particular that if one adopts the idealizational conception of science [Nowak, 1971; Nowak, 1977; Nowak, 1980], one will be able to understand, on the one hand, how one theory could be irreducible to another even though there are type-type identities among the theories' relevant predicates. Moreover, the proposal allows one to escape the charge of type-epiphenomenalism as well as Kim's challenge to non-reductive materialism.

1 A methodological interlude

There is a sense that the debate between reductionists (convinced that all science will be ultimately reduced to physics) and anti-reductionists (convinced that the thesis of reductionism is false) is deeply unsatisfying to anyone who has a deep respect for science and in particular for its ability to surprise us intellectually. This kind of philosophical legislation of what science will (or even may) do is perhaps particularly irritating with respect to psychology, which is by no accounts developed enough for us to even begin to see any overarching theories or even the direction for such, not to speak of the possibility of reducing them.

In this kind of situation, one faces a choice. One can try to find philosophical arguments for either of the sides—one can either become a reductionist or an anti-reductionist. Another option is to become an agnostic and refuse to take sides insisting on the need to look at the developments in the actual sciences before one could make the relevant judgments. This second option can take two flavors. One can become a passive agnostic and simply not get involved in the debate. But equally well, one can become an active agnostic and while refusing to take a stand one can try to contribute to the debate, in particular by multiplying the various theoretical possibilities.

This paper is written in the spirit of just such an active agnosticism. Moreover, such an attitude can help us to better understand the major change that took place on just this issue in the last century. While at the beginning of the 20th century, philosophical intuitions seemed to lie on the side of reductionism, the end of the 20th century has been taken over by a 'fashion' for anti-reductionism. The term 'fashion' is justified because the major shift in our positive attitude toward anti-reductionism has not been supported by proportionally good ways of understanding how anti-reductionism is possible (as the critics of anti-reductionism have been eager to point out, see e.g. [Kim, 1998]). This stands in stark contrast to the position of reductionism, which has been rather clearly understood (see e.g. Lewis [1966; 1972]). However, if we accept the position of active agnosticism (as well

as if we assume that the philosophical community has more or less consciously accepted such an attitude), then we can understand its predominant anti-reductionism as more than a fashion: it is precisely because reductionism is better understood while anti-reductionism is faced with numerous problems that we need to invest our energy into a better understanding of how anti-reductionism is at all possible.

2 E. Nagel's ontological reduction

We will begin with the widely accepted starting point for both reductionist and anti-reductionist conceptions, viz. E. Nagel's [1961] conception of reduction. Nagel distinguishes between homogeneous reduction (where the theories reduced use the same concepts, as he conjectured is the case in the reduction of Galileo's law of free fall to Newton's law of gravitation, for example) and heterogeneous reduction (where the theories use different conceptual schemes, as is the case in the reduction of 'qualitative thermodynamics' to molecular thermodynamics, for example). If our target is the reduction of a (developed) psychological theory to a (developed) physical or physiological theory, we should be focusing on heterogeneous reduction.[2]

E. Nagel lists numerous conditions both formal and empirical that are satisfied by the reduced and the reducing theories. They include the justification of the theories, their historical development, their systematization. One of the two central formal conditions is the requirement that the terms of the reduced theory that do not appear in the reducing theory must be expressed in terms of the reducing theory. Such relations between the terms of the respective theories are captured by the coordinating definitions or 'bridge laws'. According to the second main condition:

> (C) the experimental laws of the reduced theory are logical consequences of the theoretical postulates and coordinating definitions of the reducing theory.

It should be thus stressed that reduction so conceived is based on coordinating definitions, the so-called bridge laws, which express type-type relations (identity, or

[2] I believe that the distinction has in fact only led to problems. Homogeneous reductions, which have been regarded by Nagel as so unproblematic, have in fact given rise to one of the most notorious problems in philosophy of science, the problem of incommensurability [Kuhn, 1962/1970]. However, there is a not frequently noticed though relatively easy way out of the problem once one only gives up on the distinction. It is relatively natural to think that the concepts of the reduced are being *explicated* in terms of the concepts of the reducing theories in the cases of heterogeneous reductions. If that thought is accepted, there is but a short step to thinking that the very same thing happens in the cases of so-called homogenous reductions. This, of course, raises the problem of what explication is (I make some suggestions on how to approach it in Paprzycka [1999] and [forthcoming]), but this is a worthwhile philosophical query in any case, and the benefit of this philosophical strategy is that it reduces one philosophical problem (incommensurability) to another—broader—one (the problem of the nature of explication). (This strategy has been partially endorsed in a rather not well-known paper by James Gaa [1975].)

nomological equivalence) between concepts of the reduced and reducing theories.

3 Irreducibility I: Token-token identity

Is it possible for there to be two true theories explaining the same phenomena, where neither one reduces to the other? Davidson [1970] offers a positive answer. Reduction will be impossible if the concepts of the theories in question cannot be related to one another by means of bridge laws, i.e. if there are no type-type identities, even though there are token-token identities among mental and physical events. Mental event-tokens just are physical event-tokens. But such event-tokens can be described by means of mental concepts or by means of physical concepts. These two conceptualizations cut the nature at different joints and cannot be identified one with the other. Psychological theories will be accordingly irreducible to physiological or physical theories.

It is thus clear that it is the rejection of type-type identity relations that underlies Davidson's anti-reductionism.

The Charge of Type-Epiphenomenalism. It has been objected that Davidson's view leads to a version of epiphenomenalism, viz. type-epiphenomenalism.[3] It is worth pointing out that the view is free of classical epiphenomenalism, according to which mental events are causally mute. According to the token-token identity theory, mental event-tokens are causally efficacious in exactly the way in which physical event-tokens are since mental event-tokens are identical with physical event-tokens. In this view, however, mental events are not causally efficacious qua mental or with respect to their mental properties.

Consider the following example: John shouts 'The world is wonderful' so loud that the glass in the window breaks, and his depressed neighbour interrupts her garden work to throw some wilted flowers through the newly opened spaces (i.e. the broken window). In this case, John's shout is the cause of at least two events—of the breaking of the glass, on the one hand, and of the fact that his neighbour becomes upset, on the other. In the first case, it is clear that only physical properties of the shout are in play—its rapidly rising amplitude. In the second case, other properties are in play—semantic properties (if the depressed neighbour disagrees with the content of the shout) as well as psychological properties (if she is appalled by John's state of mind thus manifested).

This example is just an illustration of the fact that causal relations pertain not just between events (as is suggested by Davidson) but rather between events with respect to certain properties that the events exemplify. In the above case, the event of John's shouting is a cause of two different events with respect to two different

[3] F. Stoutland [1976] was the first to put forward this charge against Davidson's view. Later the objection has been presented in numerous versions by Honderich [1982], McLaughlin [1989], Kim [1993a].

properties. With respect to its rapidly rising amplitude, it causes the window glass to break; with respect to its content it upsets his already depressed neighbour.

While Davidson resists this move [1993], he has been—with good reason[4]—taken to privilege the physical properties and so understood as claiming that mental/
physical events cause other mental/physical events *qua* physical, i.e. with respect to their physical properties. But if so, then there is no room in Davidson's theory for the claim that mental/physical events can cause mental/physical events with respect to mental properties since Davidson resists the identification of mental properties with physical properties.[5]

It is worth stressing here that the problem arises because Davidson gives up on type-type identities. He could claim that events could cause other events with respect to their mental properties if mental properties could be identified with physical properties. But this is precisely what he denies, and what underlies his anti-reductionism. We are thus faced with a dilemma: if we want to accept Davidson's anti-reductionism, we are committed to rejecting type-type identities, which leads to type-epiphenomenalism; to avoid type epiphenomenalism, we would have to accept type-type identities but at the cost of rejecting anti-reductionism. It turns out that this is a false dilemma. We can accept type-type identities and find room for anti-reductionism.

4 Irreducibility II: Non-natural type-type identity

Before going on to suggest how one can find room for type-type identities and in principle irreducibility, it will pay to mention one other way of addressing the issue. Jerry Fodor [1974] has famously tried to resist the reduction based on the intuition made famous by Putnam [1967], viz. the multiple realizability of psychological concepts. In brief, Fodor argued that reduction requires that the bridge laws register nomological relations between natural kinds and that any bridge laws

[4]While Davidson resists accepting the 'x causes y with respect to property Z' idiom, he does accept the physicalistic Principle of the Nomological Character of Causality, according to which every causal relation between concrete events is subsumed under some strict law of physics. He thus opens himself up to the following argument. Take any causal relation between event-tokens, a causes b. Now, according to the Principle of the Nomological Character of Causality, there is a universal law which subsumes this causal relation, i.e. $(x)(y)(P(x) \rightarrow R(y))$, where $P(a)$ and $R(b)$. But now the critic will jump in and say that this is just what it means to say that 'a causes b with respect to property P'.

[5]There is some room open to manoeuver on Davidson's view. He could claim that the Principle of the Nomological Character of Causality does not settle it that there is *only one* law that subsumes a given cause (or even a given causal relation). This would make intelligible the distinction between a causing b in virtue of one property and a causing c in virtue of another property. But one would have to give up either Davidson's physicalism (the view that only physics is privy to universal causal laws) or his universalism (the view that causal relations are underwritten by universal causal laws) in order to be able to claim that an event can cause others in virtue of other than physical properties. (Moves in this direction have been suggested by McDowell [1985]).

between psychological concepts and physical concepts would not capture identities between natural kinds because the only way to think of psychological kinds in physical terms would be necessarily disjunctive and disjunctive kinds are not natural kinds.

Fodor's position has been challenged by Kim [1992] who argues that in the case where one 'higher' natural kind is a disjunctive kind based on some other 'lower' natural kinds, we would not thereby have 'higher' laws pertaining to the 'higher' natural kind, which would be irreducible to the 'lower' laws pertaining to the 'lower' natural kind (as Fodor thinks), but rather we would simply have to recognize that the 'higher' natural kind is not a natural kind at all. This is because any causal power that this 'higher' natural kind has are entirely derivative from those of the 'lower' natural kinds. Kim recalls a case from the history of mineralogy to support this thesis, viz. the relationship between jade, which was once thought to be a natural kind, but this view was rejected once it was discovered that jade is not a homogenous stone, but is made of two natural-kind stones, nephrite and jadeite, occurring in various combinations.

Fodor's [1997] response to Kim is to allow for the possibility of their being purely disjunctive natural kinds for which there are no empirical laws but Fodor argues that it would be simply question-begging to disallow the possibility of "higher-order" natural kinds for which there are independent empirical laws. [6] The very same point can be, of course, levelled at Fodor from Kim's perspective. We arrive at a standstill.

Kim's Dilemma. The reasoning underlying Kim's [1989] could be, however, presented in the form a dilemma to the antireductionists. If it is indeed the case that there are higher-order psychological regularities that are irreducible to physical regularities, this means that the mental properties cited in such regularities have causal powers. The causal powers of those mental properties can be thought to be either dependent on the causal powers of the physical properties or to be independent of them. If the first option holds, if the causal powers of mental properties are dependent on the causal powers of the physical properties, we could not claim that any *sui generis* psychological regularities are discovered—psychology would then be reducible to physics if only locally (this corresponds to the jade case and the position advocated here is that of reductive materialism). If the second option holds, if the causal powers of psychological properties are independent of the causal powers of physical properties, this is tantamount not to the position of non-reductive materialism but to the position of (property) dualism, which is well-known for its problems. In either case, it seems that the position of non-reductive materialism (also in Fodor's version, i.e. of non-natural type-type identity theory) is unstable—if one resists the reductive version of materialism, one seems to be

[6] Similar distinctions have been in effect proposed by Block[1997].

committed to dualism.

5 Reduction in terms of the idealizational conception of science

Although much work has been done on homogeneous reduction in idealizational terms [Krajewski, 1977; Nowakowa, 1975; Nowakowa, 1994; Paprzycka, 1990], there has not been a systematic treatment of heterogeneous reduction.

For our purposes, we will simplify the discussion by simply assuming that the coordinating definitions express identities between the concepts of the reducing and the reduced theory. Some words of justification are in order. The assumption is not meant to be even suggestive of the way things are. It is likely to turn out to be false—there are a lot of intricacies here that we are simply pushing to one side. The justification for the assumption has to do with its place in the debate between reductionism and anti-reductionism: we are making the assumption that seems to be the safeguard of reductionism, and we will be showing that—under appropriate conditions—it can lead to anti-reductionism just as well.

Since we are applying the tools of the Idealizational Conception of Science, we will need to add another assumption to Nagel's assumptions, viz. that the theories in question are simple idealizational theories.[7] In addition (for details, see [Krajewski, 1977; Nowakowa, 1975; Nowakowa, 1994; Paprzycka, 1990]), we will assume that the theories t and T investigate the same magnitude and the space of factors considered by $t$to be essential to the investigated magnitude is a subset of the space of factors considered by $T$to be essential to the investigated magnitude. The cases where a reduction involves the addition of a new essential factor have been considered in the discussion of homogeneous reduction. Here we will only consider a case where the space of essential factors is identical between the two theories.

What certainly needs to modified is condition (C). We will say that the idealizational theory t is reduced to the idealizational theory Tin case:

(C_i) T together with appropriate coordinating definitions makes it possible to derive* the idealizational law of t as well as its concretizations.

Two points are in order. First, the term 'derive*' is meant to be noncommittal as to the exact nature of the derivation. It could be that sometimes the derived* theorems are the logical consequences of the reducing theory, but they can also be approximations [Schaffner, 1967] or limiting cases thereof. We will simply not

[7] The basic terminology of the Idealizational Conception of Science necessary for our purposes is presented in the next section where appropriate schematic examples are introduced. For a detailed presentation of the view, see Nowak [1977; 1980], for the summary of further developments, see Nowak [**?**; 2000], and for various applications, see Nowak & Nowakowa [2000].

pursue this point further (and I will use 'derive' without the asterisk from now on) since this is quite a general problem for any account of reduction.

The second point, however, is crucial and it bears emphasizing. In the case of the idealizational construal of theories, there is a choice of how the derivational condition is to be construed. One could, for instance, demand that only the last concretization of the reduced theory t be derived from the reducing theory, but not that the idealizational law as well as all its concretizations be so derived. In such a case, however, it would be hard to speak of the reduction of the *theory*, one could at best speak only of the subsumption of the empirical results obtained by the theory to another. Part of what is involved in the reduction of one theory to another is not just that certain regularities that were accounted for by the reduced theory are also accounted for by the reducing theory, but also that some of the theoretical and explanatory work done by the reduced theory can be—to some extent—inherited and preserved by the reducing theory. If the subsumption of empirical results were all that mattered then we might have to look for reduction relations between contemporary scientific medicine and 'witchcraft medicine'. On the positive side, some of the classical examples of reduction do conform to (C_i). The (idealizational) ideal gas law can be derived from statistical mechanics by means of appropriate coordinating definitions, and likewise can one derive its concretizations proposed by van der Waals (see Kuipers [1982; 1985; 1990]).

6 Irreducibility III: Essential incompatibility

Intuitively, two idealizational theories can be proposed for the same domain of phenomena, which they can conceptualize in different ways. Even if the factors considered to be essential to the investigated magnitude may be identical, their essential structures will differ – what according to one theory is the most essential factor may be inessential according to the other.

Let us assume that we are considering two simple idealizational theories: theory t and theory T, whose domains are identical. The theory t investigates factor C, for which M, m_1 and m_2 are essential.[8] The *essential structure* of C is the hierarchy:

$$\begin{array}{ll} S_C: & M \\ & M, m_1 \\ & M, m_1, m_2 \end{array}$$

[8]The notion of essentiality as well as that of the degrees of essentiality are primitive in the Idealizational Conception of Science (Nowak [1977; 1980]). Intuitively, to say that factor x is essential to factor y is to say that factor x influences factor y, and the degree of x's essentiality for y expresses the extent of such influence. However, these intuitive stipulations are far from satisfactory. Moreover, given the centrality of the concept, it is one of the main research projects to try to offer explications for it. With various degrees of success, some work in this direction has been carried out by Nowak[1989], Paprzycka & Paprzycki [1992], Paprzycka [forthcoming].

where M is the principal (the most essential) factor, while m_1 and m_2 are secondary factors of diminishing essentiality.

We will assume that t is composed of the following *idealizational law*

$$(t_0)(x)[m_1(x) = 0 \wedge m_2(x) = 0 \rightarrow C(x) = g_0(M(x))],$$

where the expression '$m_1(x) = 0$' is an *idealizing assumption*, whose effect is to assume that factor m_1 does not exert any influence on factor C, and the function g_0 expresses the relation between the principal factor M and the investigated factor C on the assumption that secondary factors do not exert any influence on C–g_0 is also said to express the *regularity*.

We will assume further that the idealizational law (t_0) has been concretized first with respect to the more essential secondary factor, viz. m_1, by postulating and testing the first concretization of the idealizational law:

$$\begin{aligned}(t_1)(x)[m_1(x) \neq 0 \wedge m_2(x) = 0 \rightarrow C(x) = &\quad h_{k1}(g_0(M(x)), k_1(m_1(x)))\\ = &\quad g_1(M(x), m_1(x))],\end{aligned}$$

where the so-called *corrective function* k_1 shows how the thus far neglected factor m_1 influences the investigated magnitude, and the *directional function* h_{k1} shows how the corrective function modifies the regularity. When these two functions are superimposed, we can speak of the function g_1, which shows how factors M and m_1 affect the investigated magnitude.

We will also assume that the last concretization, which takes into account factor m_2 has been carried out:

$$\begin{aligned}(t_2)(x)[m_1(x) \neq 0 \wedge m_2(x) \neq 0 \rightarrow C(x) &= h_{k2}(g_1(M(x), m_1(x)), k_2(m_2(x))\\ &= g_2(M(x),\ m_1(x), m_2(x))],\end{aligned}$$

where k_2 is the corrective function, h_{k2} is the directional function, and g_2 expresses the dependence of factor C on all factors essential to it. The last concretization of the idealizational law is also called the *factual statement* since it no longer applies to idealized models but rather to reality. The sequence of statements (t_0), (t_1), (t_2) is called a *simple idealizational theory*.

We will assume further that theory T investigates factor D, for which factors N, n_1 and n_2 are essential. The theory T assumes the following essential structure S_D:

$$\begin{aligned}S_D: \quad &N\\ &N, n_1\\ &N, n_1, n_2\end{aligned}$$

We will assume that theory T is composed of the following idealizational law:

$$(T_0)(x)[n_1(x) = 0 \wedge n_2(x) = 0 \rightarrow D(x) = f_0(N(x))]$$

and its concretizations:

$$(T_1)(x)[n_1(x) \neq 0 \wedge n_2(x) = 0 \rightarrow D(x) = f_1(N(x), n_1(x))]$$
$$(T_2)(x)[n_1(x) \neq 0 \wedge n_2(x) \neq 0 \rightarrow D(x) = f_2(N(x), n_1(x), n_2(x))],$$

where f_0 expresses the regularity of theory T, while f_1 and f_2 express the dependence of the investigated magnitude on the appropriate essential factors.

Let us assume further that it was discovered that (a) the investigated magnitudes are identical

$$C = D,$$

and (b) that the remaining factors of theory t are likewise identifiable as factors of theory T. It is here that a crucial distinction arises. If the factors of both theories are identified in such a way that there is an isomorphism between the essential structures of the two theories, then the theories are *essentially compatible*; if there is no isomorphism between them, they are *essentially incompatible*.

Essentially Compatible Theories. In the above example, theories T and t will be essentially compatible if the following coordinating definitions are in order:

$$M = N$$
$$m_1 = n_1$$
$$m_2 = n_2.$$

In this way, the essential ordering of the factors is preserved. If so, then it is possible to derive (t_0) from (T_0), and the same holds for the respective concretizations. This structure is represented in Figure 1.

The case of essentially compatible theories exemplifies the relation between theories that corresponds to reductionist intuitions. If theories are essentially compatible, then it is possible to derive idealizational law of the reduced theory t from the idealizational law of the reducing theory T, and it is also possible to derive each concretization of the idealizational law of t from the respective concretization of the idealizational law of T. An example here would be the reduction of qualitative thermodynamics to statistical thermodynamics, where both the idealizational law (the ideal gas law) and its concretizations (van der Waals' corrections) can be derived from statistical thermodynamics (see Kuipers [1982; 1985; 1990]).

Essentially Incompatible Theories. In the case of essentially incompatible theories, the essential factors of theories t and T are related by means of coordinating definitions that do not preserve the isomorphism between the essential orderings of both theories (identifying principal factors of one theory with secondary factors of another). Let us take as a schematic example the following coordinating definitions:

$$M = n_1$$
$$m_1 = n_2$$
$$m_2 = N.$$

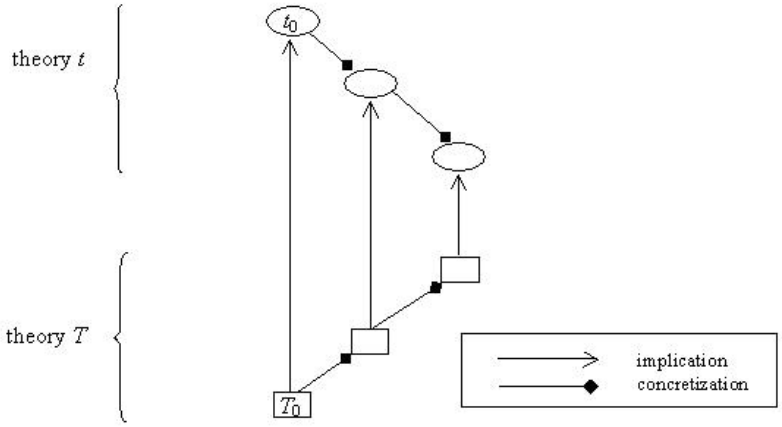

Figure 1. Essentially compatible idealizational theories

Given such coordinating definitions, it is impossible to reduce theory t to theory T (if we accept (C_i). Consider just the consequents of the idealizational laws first. According to (T_0), there are some conditions where the following dependence holds:

$$D(x) = f(N(x))$$

But given coordinating definitions, this dependence in t amounts to:

$$C(x) = k_1(m_2(x)),$$

which is but a corrective function, showing how the secondary factor m_2 influences the investigated magnitude. This is no way near the dependence necessary for bringing about the reduction of t to T, i.e. $C(x) = g(M(x))$. The more so that we also need to take into account the conditions under which the dependences are to obtain. Given the above coordinating definitions, from T we will be able to derive:

$$(T_0')(x)[M(x) = 0 \wedge m_1(x) = 0 \rightarrow C(x) = h(m_2(x))]$$
$$(T_1')(x)[M(x) \neq 0 \wedge m_1(x) = 0 \rightarrow C(x) = h'(m_2(x), M(x))]$$
$$(T_2')(x)[M(x) \neq 0 \wedge m_1(x) \neq 0 \rightarrow C(x) = h''(m_2(x), M(x), m_1(x))]$$

While it is noteworthy that we will be able to derive the factual statement of theory t, the derivation condition of reduction is not satisfied for the idealizational statements of the theory. The relationship between essentially incompatible idealizational theories is shown in Figure 2.

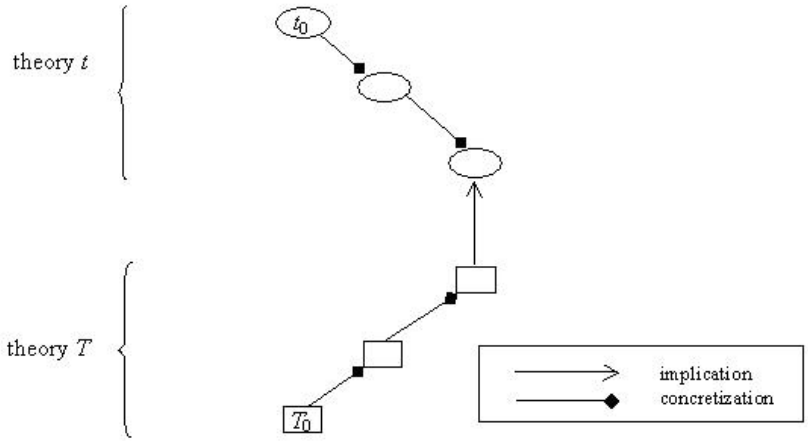

Figure 2. Essentially compatible idealizational theories

It is important to emphasize, on the one hand, that it is possible to derive the factual statements of essentially incompatible theories and, on the other, that it is impossible to derive their idealizational statements. The fact that it is impossible to derive the idealizational statements together with the idealizational law, which registers the regularity obtaining between the investigated and the principal factor, is intrinsically related to the fact that it is impossible to preserve the explanatory structure of the old theory in the new theory. According to Nowak [1977; 1980], explanation consists in showing first how the principal factor affects the investigated magnitude in the absence of secondary factors, and then slowly modifying that relationship to take into account the secondary factors so as to yield the empirical relationship exemplified in a given case. Given that the idealizational statements of the theories will not be preserved, the peak of the explanatory process will not be preserved either. This tallies nicely with the thought that the failure of reduction goes hand in hand with the failure of explanation.

However, it is also important to emphasize that it is possible to derive factual statements of essentially incompatible theories. This means that the relation between the theories is non-accidental. They allow for there to be a convergence between the theories on detailed and relatively particularized generalizations (what Nancy Cartwright [1983] has called phenomenological laws). In other words, these theories explain the same phenomena but in different ways.

7 Objections reconsidered

The Threat of Type-Epiphenomenalism. We will remember that the objection begins with the thesis that all causal relations among events obtain in virtue of their physical properties. The objection arises when—in order to rescue anti-reductionism—one rejects the type-type identity theory and is consequently unable to uphold the thesis that causal relations obtain in virtue of their mental properties. The proposal presented above shows how it is possible to accept non-reductionism without denying type-type identities. We can thus reject what was blocking the intuitive view that causal relations obtain not only in virtue of the physical but also in virtue of the mental properties. Essential incompatibility between idealizational theories is based on type-type identities between the factors of the theories in question without, however, preserving their essential orderings. This guarantees differences in the idealizational statements of both theories, preserving not only a token-token but also a type-type convergence on what but not how the theories explain.

The Threat of Dualism. We will remember that Kim's dilemma starts with a choice: either one chooses the view that mental causal powers are dependent on physical causal powers, which seems to commit one to reductive materialism, or one chooses the view that mental causal powers are independent of physical causal powers, which seems to commit one to dualism. There seems to be no room for a non-reductive version of materialism in this scenario.

Again, our discussion shows that there is a third way. We can accept for the sake of the argument that the causal powers of mental properties depend on the causal powers of physical properties. This does not yet prejudge the fact that psychological theories will be reducible to physical theories because the causal powers of properties do not determine the essential ordering of factors.

8 Conclusion

We have seen how the debate between reductionism and anti-reductionism can be enriched by taking into account developments in the philosophy of science. Using the framework of the Idealizational Conception of Science, I have shown that it is possible to uphold the non-reductionist view while at the same time avoiding at least some of the objections that have been leveled against it. I have argued that taking the idealizational structure into account allows one to drive a wedge between type-type identity and reductionism. One way to claim that one theory does not reduce to another is to claim that although there is an identity among tokens, there is no identity among types, and so that there is no reduction of any generalizations. The other way to claim that one theory does not reduce to another is to claim that although there are type-identities among the factors of the theories, there are still no correspondences between the ways in which these factors figure

into the hierarchy of laws of the theories in question. And while there may be correlations between some distant concretizations at the bottom of the hierarchies, this does not amount to the reduction of one theory to the other, since there is no correlation between the most general laws at the top of the hierarchies. In short, it is no longer imperative to postulate only token-token identities to save us from reductionism.

The conclusion of this paper is only apparently anti-reductionist. We have seen the conditions (of essential incompatibility of theories) under which the reduction of two idealizational theories would be impossible. Whether such a relation takes place among any theories—and psychology and neurophysiology in particular—is and remains an open question. The task of philosophy is to understand the very nature of reduction and autonomy. Whether reductionism or anti-reductionism are the correct views to hold will, however, be decided by those who are best equipped to make this decision—-by scientists in the near and, mostly likely, the distant future.

Acknowledgements

I have benefited greatly from comments and suggestions made by a number of people. I want to mention in particular: Jeremy Butterfield, Theo Kuipers, Krzysztof Łastowski, Leszek Nowak, Włodzimierz Rabinowicz and Marcel Weber.

BIBLIOGRAPHY

[Block, 1997] N. Block. Anti-Reductionism Slaps Back. *Philosophical Perspectives*, **11**, 107–132, 1997.

[Brzezinski *et al.*, 1990] J. Brzeziński, F. Coniglione, T.A.F. Kuipers and L. Nowak, eds. *Idealization I: General Problems*. Amsterdam: Rodopi, 1990.

[Cartwright, 1983] N. Cartwright. *How the Laws of Physics Lie*. Oxford: Clarendon Press, 1983.

[Cartwright, 1989] N. Cartwright. *Capacities and their Measurement*. Oxford: Oxford University Press, 1989.

[Davdison, 1970] D. Davidson. Mental Events. Reprinted in *Essays on Actions and Events* , pp. 207–228. Oxford: Clarendon Press, 1970.

[Davidson, 1993] D. Davidson. Thinking Causes. In J. Heil and A. Mele, eds. *Mental Causation*, pp. 3–18. Oxford: Clarendon Press, 1993.

[Fodor, 1974] J. Fodor. Special Sciences, or the Disunity of Science as a Working Hypothesis. Reprinted in N. Block, *Readings in Philosophy of Psychology*, pp. 120–133. Cambridge, MA: Harvard University Press, 1980.

[Fodor, 1997] J. Fodor. Special Sciences: Still Autonomous after All these Years. *Philosophical Perspectives*, **11**, 149–163, 1997.

[Gaa, 1975] J. Gaa. The Replacement of Scientific Theories: Reduction and Explication. *Philosophy of Science*, **42**, 349–372, 1975.

[Honderich, 1982] T. Honderich. The Argument for Anomalous Monism. *Analysis*, **42**, 59–64, 1982.

[Kim, 1989] J. Kim. The Myth of Nonreductive Materialism. Reprinted in *Supervenience and Mind: Selected Philosophical Essays*, pp. 265–284. Cambridge: Cambridge University Press, 1993.

[Kim, 1993a] J. Kim. Can Supervenience and 'Non-Strict Laws' Save Anomalous Monism? In J. Heil, and A. Mele, eds. *Mental Causation*, pp. 19–26. Oxford: Clarendon Press, 1993.

[Kim, 1992] J. Kim. Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research*, **52**, 1–26, 1992.

[Kim, 1993b] J. Kim. *Supervenience and Mind. Selected Philosophical Essays*. Cambridge: Cambridge University Press, 1993.
[Kim, 1998] J. Kim. *Mind in a Physical World. An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: The MIT Press, 1998.
[Krajewski, 1977] W. Krajewski. *The Principle of Correspondence and the Growth of Knowledge*. Dordrecht-Boston: Reidel, 1977.
[Kuhn, 1962/1970] T. Kuhn. *The Structure of Scientific Revolutions*, 2nd edition, Chicago: The University of Chicago Press, 1970.
[Kuipers, 1982] T. A. F. Kuipers. The Reduction of the Phenomenological to Kinetic Thermostatics. *Philosophy of Science*, **49**, 107–119, 1982.
[Kuipers, 1985] T. A. F. Kuipers. The Paradigm of Concretization: The Law of van der Waals. In J. Brzeziń ski, ed., *Consciousness: Methodological and Psychological Approaches: Poznań Studies in the Philosophy of the Sciences and the Humanities*, vol. 8, pp. 185–199. Amsterdam: Rodopi, 1985.
[Kuipers, 1990] T. A. F. Kuipers. Reduction of Laws and Concepts. In [Brzezinski *et al.*, 1990, pp. 241–276].
[LePore and McLaughlin, 1985] E. LePore and B. P. McLaughlin, eds. *Actions and Events. Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell, 1985.
[Lewis, 1966] D. Lewis. An Argument for the Identity Theory. *Journal of Philosophy*, **63**, 17–25, 1966.
[Lewis, 1972] D. Lewis. Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy*, **50**, 249–258, 1972.
[McDowell, 1985] J. McDowell. Functionalism and Anomalous Monism. In [LePore and McLaughlin, 1985, pp. 387–398].
[McLaughlin, 1989] B. P. McLaughlin. Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical. *Philosophical Perspectives*, **3**, 109–136, 1989.
[Nagel, 1961] E. Nagel. *The Structure of Science*. New York: Harcourt, Brace and World, 1961.
[Nowak, 1971] L. Nowak. *U podstaw marksowskiej metodologii nauk*, Warszawa: PWN, 1971.
[Nowak, 1977] L. Nowak. *Wstęp do idealizacyjnej teorii nauki*. Warszawa: PWN, 1977.
[Nowak, 1980] L. Nowak. *The Structure of Idealization*. Dordrecht/Boston: Reidel, 1980.
[Nowak, 1989] L. Nowak. On the (Idealizational) Structure of Economic Theories. *Erkenntnis*, **30**, 225–246, 1989.
[Nowak, 2000] L. Nowak. The Idealizational Approach to Science: A New Survey. In [Nowak and Nowakowa, 2000].
[Nowak and Nowakowa, 2000] L. Nowak and I. Nowakowa. *The Richness of Idealization* (*Poznań Studies in the Philosophy of the Sciences and the Humanities*, vol. 69. Amsterdam/Atlanta: Rodopi, 2000.
[Nowakowa, 1975] I. Nowakowa. *Dialektyczna korespondencja a rozwój nauki*. Warszawa-Poznań: PWN, 1975.
[Nowakowa, 1994] I. Nowakowa. *The Dynamics of Idealizations*. Amsterdam-Atlanta: Rodopi, 1994.
[Paprzycka, 1990] K. Paprzycka. Reduction and Correspondence in the Idealizational Approach to Science. In [Brzezinski *et al.*, 1990, pp. 277–286].
[Paprzycka, 1999] K. Paprzycka. Socrates Meets Carnap: Explication in the Theaetetus. *Logical Analysis and History of Philosophy*, **2**, 87–108, 1999.
[Paprzycka, forthcoming] K. Paprzycka. *Oblicza anty-redukcjonizmu (Faces of Anti-Reductionism), forthcoming.*
[Paprzycka and Paprzycki, 1992] K. Paprzycka and M. Paprzycki. Accuracy, Essentiality, Idealization. In J. Brzeziński and L. Nowak, *Idealization III: Approximation and Truth* (*Poznań Studies in the Philosophy of the Sciences and the Humanities*, vol. 25). Amsterdam/Atlanta, GA: Rodopi, 1992.
[Putnam, 1967] H. Putnam. The Nature of Mental States. Reprinted in *Mind, Language and Reality: Philosophical Papers*, vol. II, pp. 429–440. Cambridge: Cambridge University Press, 1975.
[Salmon, 1984] W. C. Salmon. *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press, 1984.

[Salmon, 1989] W. C. Salmon. Four Decades of Scientific Explanation. In P. Kitcher and W.C. Salmon, eds, *Scientific Explanation. Minnesota Studies in the Philosophy of Science, Vol. XIII*, pp. 3–219. Minneapolis: University of Minnesota Press, 1989.

[Schaffner, 1967] K. F. Schaffner. Approaches to Reduction. *Philosophy of Science*, **34**, 137–147, 1967.

[Soutland, 1976] F. Stoutland. The Causation of Behavior. In J. Hintikka, ed., *Essays on Wittgenstein in Honor of G.H. von Wright: Acta Philosophica Fennica*, **28**, 286–325, Amsterdam: North-Holland, 1976.

[Stoutland, 1985] F. Stoutland. Davidson on Intentional Behaviour. In [LePore and McLaughlin, 1985, pp. 44–59].

Katarzyna Paprzycka
Instytut Filozofii
Szkoła Wyższa Psychologii Społecznej (SWPS)
u. Chodakowska 19/31
03-815 Warszawa, Poland
Email: Katarzyna.Paprzycka@swps.edu.pl